

This is a repository copy of *Detecting joint attention events in mother-infant dyads : Sharing looks cannot be reliably identified by naïve third-party observers*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/176768/>

Version: Published Version

---

**Article:**

Graham, Kirsty E, Buryen-Weitzel, Joanna C, Lahiff, Nicole J et al. (2 more authors) (2021) Detecting joint attention events in mother-infant dyads : Sharing looks cannot be reliably identified by naïve third-party observers. PLoS ONE. e0255241. ISSN 1932-6203

<https://doi.org/10.1371/journal.pone.0255241>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

RESEARCH ARTICLE

# Detecting joint attention events in mother-infant dyads: Sharing looks cannot be reliably identified by naïve third-party observers

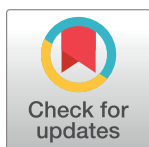
Kirsty E. Graham<sup>1,2</sup>\*, Joanna C. Buryn-Weitzel<sup>1‡</sup>, Nicole J. Lahiff<sup>1‡</sup>, Claudia Wilke<sup>1‡</sup>, Katie E. Slocombe<sup>1</sup>

**1** Department of Psychology, University of York, York, United Kingdom, **2** School of Psychology & Neuroscience, University of St Andrews, St Andrews, United Kingdom

\* These authors contributed equally to this work.

‡ JCB-W, NJL and CW also contributed equally to this work.

\* [keg4@st-andrews.ac.uk](mailto:keg4@st-andrews.ac.uk)



## Abstract

Joint attention, or sharing attention with another individual about an object or event, is a critical behaviour that emerges in pre-linguistic infants and predicts later language abilities. Given its importance, it is perhaps surprising that there is no consensus on how to measure joint attention in prelinguistic infants. A rigorous definition proposed by Saposova & Carpenter (2019) requires the infant and partner to gaze alternate between an object and each other (coordination of attention) and exchange communicative signals (explicit acknowledgement of jointly sharing attention). However, Hobson and Hobson (2007) proposed that the quality of gaze between individuals is, in itself, a sufficient communicative signal that demonstrates sharing of attention. They proposed that observers can reliably distinguish “sharing”, “checking”, and “orienting” looks, but the empirical basis for this claim is limited as their study focussed on two raters examining looks from 11-year-old children. Here, we analysed categorisations made by 32 naïve raters of 60 infant looks to their mothers, to examine whether they could be reliably distinguished according to Hobson and Hobson’s definitions. Raters had overall low agreement and only in 3 out of 26 cases did a significant majority of the raters agree with the judgement of the mother who had received the look. For the looks that raters did agree on at above chance levels, look duration and the overall communication rate of the mother were identified as cues that raters may have relied upon. In our experiment, naïve third party observers could not reliably determine the type of look infants gave to their mothers, which indicates that subjective judgements of types of look should not be used to identify mutual awareness of sharing attention in infants. Instead, we advocate the use of objective behaviour measurement to infer that interactants know they are ‘jointly’ attending to an object or event, and believe this will be a crucial step in understanding the ontogenetic and evolutionary origins of joint attention.

## OPEN ACCESS

**Citation:** Graham KE, Buryn-Weitzel JC, Lahiff NJ, Wilke C, Slocombe KE (2021) Detecting joint attention events in mother-infant dyads: Sharing looks cannot be reliably identified by naïve third-party observers. PLoS ONE 16(7): e0255241. <https://doi.org/10.1371/journal.pone.0255241>

**Editor:** Ewa Pisula, University of Warsaw, POLAND

**Received:** February 24, 2021

**Accepted:** July 12, 2021

**Published:** July 23, 2021

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0255241>

**Copyright:** © 2021 Graham et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its [Supporting Information](#) files.

**Funding:** This study was funded by an European Research Council Consolidator grant (ERC\_CoG

2016\_724608) awarded to KS (<https://erc.europa.eu/funding/consolidator-grants>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

The emergence of joint attention, the ability of two individuals to share attention about an external object or event, is considered a crucial step in child development [1]. Joint attention is thought to emerge in infancy at around 9–12 months [1–3], and may be important in language acquisition, with children connecting words with objects to which they and another individual are jointly attending [1, 4, 5]. Some have also argued that engaging in joint attention lays the foundation for shared intentionality, which in turn underpins human cooperation [6]. Given the links between engaging in joint attention and both language and cooperation, it has been suggested that joint attention might represent a key species-difference between humans and other great apes [7]. However, the extent to which non-humans engage in joint attention is fiercely debated, with some studies suggesting non-human great apes have joint attention skills [8–10] and some studies finding no evidence for engagement in joint attention events [3]. This debate is currently hampered by the lack of a consistent and clear definition of joint attention events that can be applied comparatively across species.

Given the developmental and evolutionary importance of joint attention, it is perhaps surprising that there is a lack of consensus in the literature as to both what constitutes joint attention and how to operationalize definitions. Most studies focus on joint attention skills, and examine an individual's ability to initiate joint attention with a partner or respond to joint attention cues, such as the gaze or pointing gestures of a partner. Joint attention skills are necessary but not sufficient for engaging in joint attention events, which also require an individual to have the motivation and opportunity to use those skills with willing partners. What conceptually constitutes a full joint attention event, where at least two people actually engage in joint attention, remains diverse. In the literature, a 'Joint attention event' can refer to behaviours as varied as spectators in a stadium all watching the same football match, and a child following the gaze of their mother to jointly attend to a butterfly and exchanging emotional reactions to the butterfly, making the jointness of the interaction manifest [11]. Similarly, when it comes to operationalizing joint attention events, a diverse array of approaches also exists (Table 1). In order to better understand this variation, we conducted a systematic search for operational definitions of joint attention events published in English in journal articles from 2000–2018 by inserting 'Joint Attention' into Science Direct, and then manually checking to remove papers that do not define and measure joint attention events, or that only define initiating joint attention skills (IJA) or responding to joint attention skills (RJA). We also excluded studies that did not examine human-human interactions, e.g. human-video, human-robot, nonhuman-human etc. (Table 1).

These definitions of joint attention represent a wide range of behaviours and interactions with varying degrees of 'jointness'. At the simplest end, it is sufficient for two individuals to both look at the same object (also called 'parallel attention' [39]), or for two individuals to look at one another; and at the more complex end, individuals must coordinate their attention between an object or event and one another and also communicate with one another about it (Table 1). The problem inherent in the simpler operational definitions of joint attention events is that such behavioural patterns could be explained by (a) properties of the event itself (a salient event, such as a loud noise, may draw the attention of multiple individuals, but they might not be aware of others also orienting to the same event) or (b) by the desire to check or monitor the behaviour of another (an event may be of interest but so is monitoring the behaviour of another individual, resulting in gaze alternation between these competing loci of attention). Whilst monitoring or checking the behaviour of another individual might ultimately have sharing motives, one might also engage in this behaviour for more individualistic or competitive reasons (e.g. an infant may want to check the mother is remaining close-by or an

**Table 1. A list of operationalized definitions of joint attention events used in papers between 2000 and 2018 that met criteria for our literature review; references [12–104].**

Operationalized Definition for Joint Attention events	Studies that use this definition
Attend to partner's actions	Gonsiorowski et al. 2016
Both attend to an object/event at the same time ( <i>no mutual gaze required</i> )	Arens, Cress & Marvin 2005*; Cress, Arens, & Zajicek 2007*; Craig et al 2004; Hughes & Allen 2013; Kompatsiari et al 2018; Macdonald & Tatler 2018; Najnin & Banerjee 2018; Povis & Crowley 2015; Richardson et al 2007; Sung & Hsu 2009; Walberg & Craig-Unkefer 2010; Yont, Snow & Vernon-Feagans 2003; Yu & Smith 2013; Yu & Smith 2016; Yu & Smith 2017
Mutual Gaze ( <i>no attention to object required</i> )	Charman et al. 2000*; Hobson & Hobson 2007**; Morgan, Maybery & Durkin 2003; Sanefuji et al. 2009
Gaze switch: object-partner <b>OR</b> partner-object	Aarne & Tallberg 2010; Arens, Cress & Marvin 2005*; Charman et al. 2000*; Charman et al. 2003; Cress, Arens, & Zajicek 2007*; Clifford & Dissanayake 2009; Vismara et al. 2018; Yu & Ballard 2007
Gaze switch + partner communication ( <i>i.e. individual switches gaze to an object while partner is communicating about that object</i> )	Matatyaho & Gogate 2008
Gaze alternation: object-partner-object <b>OR</b> partner-object-partner	Arens, Cress & Marvin 2005*; Clifford & Dissanayake 2008; Cress, Arens, & Zajicek 2007*; Spector & Charlop 2018; Striano & Bertin 2005
Coordinates attention between partner and object/event ( <i>this involves looking at both the partner and the object/event but does not specify order, frequency, or duration</i> )	Aldrich et al 2015; Arnold et al. 2000; Brune & Woodward 2007; Cleveland & Striano 2007; Conboy et al. 2015; Gaffan et al. 2010; Gulsrud et al. 2007; Grossmann & Johnson 2010; Johnson et al. 2008; Legerstee, Markova & Fisher 2007; Neerinx et al 2014; Nelson, Adamson, & Bakeman 2008; Nordahl-Hansen et al. 2016; Saxon et al. 2000; Striano & Bertin 2005; Striano et al. 2009; Suma et al. 2016; Trautman & Rosenthal 2006; Yazbek & D'Entremont 2006
Coordinates attention "for the purpose of sharing"	Ingersoll & Schreibman 2006
Coordinates attention between partner and object/event + acknowledges partner's participation (quality of looks may be used to judge acknowledgement of partner's participation)	Adamson et al. 2009; Adamson et al. 2010; Bigelow 2003; Childers, Vaughan, & Burquest 2007; Gauthier et al. 2011; Gulsrud et al. 2016; Hahn et al. 2016; Kasari et al. 2010; Larkin et al. 2015; Markus et al. 2000; Mastin & Vogt 2016; Rice et al. 2016; Schechter et al. 2010; Schertz & Odom 2004; Schertz et al. 2018; Smith et al. 2009
Early Sociocognitive Battery (ESB): alternating gaze between partner and object	Roy & Chiat 2014
Attends to an object/event + [gazes at partner <b>AND/OR</b> communicates with partner]	Adamson, Deckner, & Bakeman 2010; Allely et al. 2013; Benigno & Farrar 2012; Benigno et al. 2007; Bigelow, Maclean & Proctor 2004; Bigelow et al. 2010; Christidou 2018; Depowski et al 2015; de la Ossa & Gauvain 2001; Henderson & Jennings 2003; Ine, Heleen & Bea 2011; Morales et al. 2005; Nowakowski, Tasker & Schmidt 2012; Nowakowski et al. 2009; Nowakowski et al. 2011; Pierce et al. 2015; Shire et al. 2016; Slaughter et al. 2008; Skarabela 2007; Skarabela & Allen 2010; Tasker & Schmidt 2008; Tasker, Nowakowski, & Schmidt 2010; Yoon et al. 2014; Zampini, Salvi & D'Odorico 2015
Coordinates attention between partner and object/event + acknowledges partner's participation + communication with partner	Chiang et al. 2016; Ketelaar et al. 2012; Landry et al. 2008

\*Papers that use more than one definition within the paper

\*\*Papers that specify types of look

<https://doi.org/10.1371/journal.pone.0255241.t001>

individual may want to check that another individual is not taking a valuable food item). It is also important to note that simpler forms of ‘joint attention’ such as the simultaneous monitoring of an object or event by multiple observers emerge earlier in development (common in 6-month olds (39)) than more complex forms of joint attention, supporting the idea that different cognitive processes may underlie some behaviours currently labelled as ‘joint attention’ in the literature.

To try and capture the ‘jointness’ of a joint attention event, more recent operational definitions seek to identify behavioural markers of mutual awareness that the interaction partners are attending to the event together and are thus sharing attention. One such behavioural marker is communication, which is a necessary element of the ‘more complex’ operational definitions in Table 1. Whilst facial expressions, such as smiles or looks of concern, and vocalisations are forms of communication that can be objectively and reliably coded from interaction partners, it has also been suggested that the quality of looks exchanged during mutual gaze can be sufficient to communicate awareness of the shared nature of the engagement (47). This suggestion is implicit in the definitions that include “acknowledges partner’s participation” (Table 1; e.g. (13), p. 19), which is often assessed by mutual gaze although there is often little description given as to what it is about the gaze that acknowledges the partner’s participation. Supporting this argument, and providing further detail on how to assess gaze, Hobson and Hobson assert that observers can reliably identify ‘sharing looks’ in children and distinguish them from checking or orienting looks [47]. In this study, two raters classified the looks of 11-year-old children with and without autism towards an experimenter into three categories according to the definitions detailed in Table 2, and significant agreement between the two

**Table 2. Definitions of types of look from Hobson & Hobson, 2007 and the summaries of these definitions provided to participants in the current study.**

Type of Look	Definitions from Hobson & Hobson (2007)	Definitions adapted for participant instructions in the current study
Sharing	“Sharing looks” were defined as those looks directed to the tester that could be seen to express a participant sharing experience through interpersonal contact with the tester. They involved a deep gaze that conveyed personal involvement . . . in contrast to checking looks that involved glances “at” the tester.	<b>Sharing look:</b> The child looked into the mother’s eyes and the look conveys personal involvement and emotional contact. The child’s look acknowledges the mother’s gaze (reciprocal)
	• <i>Reciprocal:</i> The participant appears to register that the tester is also looking to the participant.	
	• <i>Deep:</i> The look is <i>into</i> the eyes of the tester.	
	• <i>Personal:</i> The look manifests affective contact with the tester.	
Checking	“Checking looks” were defined as those looks toward the tester that were used to assess or check out either the situation or the tester’s response.	<b>Checking look:</b> The child looked at the mother’s face and eyes, likely to monitor or check on the mother’s presence and behaviour, maybe to help the child assess the situation. The child’s look does not acknowledge the mother’s gaze (not reciprocal).
	• <i>Nonreciprocal:</i> The participant appears to look to the tester without registering the tester’s look to the participant.	
	• <i>Superficial:</i> The look is at the eyes of the tester.	
	• <i>Impersonal:</i> The look appears to have the goal of monitoring the tester’s actions.	
Orienting	“Orientating looks” were those that appeared to occur in direct response to an action, sound, or movement on the part of the tester.	<b>Orienting look:</b> The child looked at the mother as a direct response to an action, sound (including language) or movement the mother made (the mother may or may not have been intending to gain their attention)

<https://doi.org/10.1371/journal.pone.0255241.t002>

raters was found. It is unclear, however, whether the looks of preverbal infants can be reliably classified in the same way and it is also important to test this with a greater number of raters.

Here, we aim to robustly test whether raters can accurately judge the type of look given by an infant to their mother as sharing, checking or orienting looks. This will enable us to evaluate whether the type of look should be considered a valid form of communication upon which judgements can be made about joint attention events occurring between preverbal infants and their mothers. We asked a group of 32 naïve raters to code infant looks to their mothers during free play according to the Hobson & Hobson definitions (Table 2). Raters judged the type of look in 63 short video clips cut 3 seconds before and after the infant looks at the mother, and also gave a rating of confidence in their judgement. For 33 videos, we had the mother's judgement as to the type of look her infant gave to her. We expected that if the types of looks could be reliably identified by naïve observers then we would obtain high levels of agreement between (i) our 32 raters, and (ii) our raters and the mothers. We aimed to explore whether self-rated confidence was associated with high levels of agreement with other raters. For looks where there was strong agreement between raters, we then performed analysis of the video to try and ascertain which behavioural cues the raters may have been basing their judgements on.

## Methods

### Ethics statement

The University of York, Department of Psychology, Departmental Ethics Committee approved the study (ID number 605). Written consent was obtained from all participants.

### Participants

We recruited 32 women in York, UK, including 14 mothers and 18 non-mothers. This allowed us to test whether experience raising infants affected agreement on the categorisation of infant looks. The average age for the mother group was 45.21 years ( $SD \pm 8.32$ ) and for the non-mother group was 32 years ( $SD \pm 8.86$ ). An independent-samples t-test revealed that the mother group were significantly older than the non-mother group ( $t(29) = 4.2, p < 0.001$ ).

### Stimuli and experimental design

Stimuli consisted of videos that were cut into short clips (mean = 8.11s,  $SD = 1.70$ s) that started 3s before the child looked at their mother and ended 3s after the child stopped looking at their mother, unless there was another look within that 3s in which case it was cut immediately before or after the second look. This time window was chosen to give raters an opportunity to observe part of the interaction in which the look occurred, but without providing broader context and behaviours, as we wanted raters to judge the quality of the look, not the general intent of the infant or mother. We presented participants with two sets of videos: "Set 1" contained 33 videos filmed in 2017, and "Set 2" contained 30 videos filmed in 2009–2010. For Set 1, the experimenters had filmed 7 mothers and infants living in or near York, UK, playing (age: mean = 15.75 months, min = 8 months, max = 26 months). We had interviewed the mothers immediately afterwards, showing them the videos and asking them to state what type of look they thought the infant had given them, using the same criteria as participants received (Table 2), with the exception that we also allowed mothers to also say if they didn't know ("Unknown"). This meant that the mothers were relatively confident in the sharing, checking and orienting look categorisations they provided. We selected groups of three videos for each mother-infant dyad, containing one of each type of look where possible (as judged by the mother). Only four looks in Set 1 were classified as "Unknown" by the mother. Three mother-



infant dyads provided one group of three looks and four dyads provided two groups of three looks. The Set 2 videos were taken from 8 mothers and infants living in or near York, UK (age: mean = 11.125 months, min = 11 months, max = 12 months) and we had no judgements from these mothers on the type of look the infant produced. We selected groups of three high quality videos for each infant. Six mother-infant dyads provided one group of three looks and two dyads provided two groups of three looks.

The experiment was presented to participants using an online platform (Gorilla.sc), which was set to “private” so that only the experimenters had access. Video dimensions for Set 1 were 16:9 with size of 21.3x12cm; Set 2 were 4:3 with 16x12cm. Within both sets of videos, the videos were shown in blocks of three that contained videos from the same mother and infant dyad. Videos were randomised within the blocks, and the order of blocks was randomised across the experiment. We also counterbalanced the order of the response buttons (three types of look) across participants. There were six possible orders for the three response buttons (sharing, checking, orienting) and participants were randomly assigned to one of these six orders. Participants were given a printed sheet of the definitions of the three types of looks (Table 2) to refer to throughout the experiment and we ensured that the order of definitions on the sheet matched the order of response buttons in the experiment.

## Procedure

The experimenter started the experiment for the participant, who was seated in a quiet room in front of a laptop computer (HP Elitebook, 14” display, 64-bit operating system). The participant was first provided with instructions on the screen, which informed them that the videos would play only once and if they saw the look they should choose one of the three types of looks (sharing, checking, orienting) even if they were uncertain, but if they did not see the look from the infant they should select the ‘did not see look’ button. We wanted to both standardise the number of times participants could view a video and to reduce the effects of fatigue on performance. Therefore, the video played only once in each trial. When the participant had selected the type of look, on the next screen they were asked to rate how confident they were in their judgement on a scale of 0–10 (0 = “Not at all confident”; 10 = “Extremely confident”) using a continuous slider. The confidence marker’s starting position on all trials was 1. There was no time limit on the duration participants spent making their judgement. Once participants had completed the confidence rating, they pressed a button that read “Next” to start the next trial, so the experiment progressed at a pace chosen by the participant.

## Video coding

To assess the possible cues that raters may have been relying on to make their judgements, we measured the duration of infant looking to mother, presence and duration of mutual gaze between infant and mother, and infant or mother communication (frequency of vocalisations, gestures or salient facial expressions during (i) the infant’s look to the mother and (ii) the whole video clip). Videos were coded using The Observer XT 14 software, and the following measures were extracted: infant looking direction (mother’s face, elsewhere), mother looking direction (infant’s face, elsewhere), and infant and mother facial expressions, gestures, and vocal communication (including vocalisations and language). All changes in the looking direction category, no matter how short were coded. The three types of communication were coded as a new instance when they changed to a new signal within the same communication type (e.g. from a smile to a frown) or when there was a gap of >1s between signals (e.g. talking, 1.2s pause, talking). Instances of each communication type were considered separately, so that a simultaneous smile and talking represent one facial expression and one vocal communication.

To assess the reliability of the video coding, KG coded all videos and CW coded 6 videos from Set 1 and 6 videos from Set 2 (20% of all videos). Using the reliability function in Observer software, we ascertained the two video coders had high levels of agreement: duration of (i) looks from the infant to their mother's face ( $\kappa = 0.88$ ) and (ii) mutual gaze between infant and mother (occurred in 4/12 videos;  $\kappa = 0.98$ ); Number of infant facial expressions, gestures, and vocalizations ( $\kappa = 0.78$ ); and number of mother facial expressions, gestures, and vocalizations ( $\kappa = 0.76$ ).

## Data analysis

It was necessary to make some exclusions from the data set prior to analysis: There were two videos where  $>3$  participants did not see the look ( $n = 16, 20$ ) and we excluded these videos entirely from analyses. There was also one video that was accidentally duplicated, and we removed the second appearance of this video from analysis. These exclusions left 30 videos for analysis in Set 1 and 30 videos for analysis in Set 2.

As the number of instances of communication produced by infants and mothers in these short video clips were low, we extracted the following measures from the coded videos: (i) during the infant's look to the mother we considered the presence/absence of each type of communication (facial, gestural, vocal) for both the infant and mother and (ii) during the whole video clip we considered the total number of communication events (sum of all vocal, gestural and facial communication events coded) produced by (a) the mother and (b) the infant.

All analyses were conducted in R 3.5.3 [105], and we used the packages "irr" for Fleiss' kappa [106], "lme4" with "lmerTest" for GLMMs [107, 108], and "ggplot2" with "plyr" for plotting [109, 110]. Descriptions of each model are included in the relevant results sections.

## Results

### Overall agreement among naïve raters

First, we calculated Fleiss' kappa to check overall inter-rater reliability for both sets of videos [111]. For Set 1 with 30 videos and 32 raters, Fleiss' kappa was 0.157 ( $z = 28.4$ ). For Set 2 with 30 videos and 32 raters, Fleiss' kappa was 0.228 ( $z = 40.1$ ). These are both low rates of overall agreement [112], suggesting that in general, participants did not agree on types of look.

We asked our 32 raters whether they had children ( $n = 14$ ) or did not have children ( $n = 18$ ), as we thought that experience with young children might affect their judgements. For Set 1, raters who were mothers had a kappa of 0.184 ( $z = 14.1$ ), and raters who were not mothers had a kappa of 0.155 ( $z = 15.6$ ). For Set 2, raters who were mothers had a kappa of 0.234 ( $z = 17.7$ ), and raters who were not mothers had a kappa of 0.226 ( $z = 22$ ). Descriptively, mothers and non-mothers seem to agree on the type of looks at comparable rates, so for all further analyses we analysed mothers and non-mothers together.

### Rater agreement on specific videos

Next, we tested whether a higher proportion of participants agreed on the type of look the infant gave in each video clip than expected by chance (0.33). A binomial test (0.33) was conducted for the type of look chosen by the highest number of raters for each video clip, to see if a higher proportion of raters than expected by chance chose that type of look. As we conducted 30 binomial tests on Set 1 and Set 2 of the videos we bonferroni-adjusted the significance threshold to 0.0016 ( $.05 / 30$  binomial tests) meaning that 20 or more raters had to be in agreement on a specific video in Set 1 or Set 2 to be significantly above the level expected by chance.



For Set 1, out of 30 videos, there were 9 videos where the raters agreed on one type of look significantly more than chance (high agreement, see Table 3). Of these videos, five of the looks were rated as “sharing looks”, two were rated as a “checking look”, and two were rated as an “orienting look” by most raters. When we compared this to how the mothers rated their own videos, the mothers agreed with only 2/5 of the sharing looks (the other three were “unknown”, “checking”, and “orienting” according to the mothers); 1/2 of the checking looks (the other was rated “sharing” by the mother); and 0/2 of the two orienting looks (rated “unknown” and “checking” by the mother; S1 Table). This means that of the 26 looks mothers had classified as sharing, checking or orienting (excluding the 4 looks mothers rated as ‘unknown’), in only 3 of these cases did a significant majority of raters agree with the categorisation of the mother.

For Set 2, out of 30 videos, there were 14 videos where the raters agreed on one look significantly more than chance (high agreement, see Table 3). Of these videos, six of the looks were rated as “sharing looks”, seven were rated as “checking looks”, and one “orienting look” by most raters.

We also tested whether raters were more or less confident for high agreement looks (as determined by the binomial tests;  $n = 23$ ) or low agreement looks (where the two highest look types had a difference of  $\leq 3$  responses;  $n = 16$ ). On a scale of 1 to 10, the mean confidence for high agreement looks was 6.19 (SE = 0.21) and for low agreement looks was 5.64 (SE = 0.20), and overall participants use a limited range of the scale (min 4.8, max 7.8). A paired t-test revealed that raters were significantly more confident about high agreement looks than low agreement looks ( $t(31) = 6.72$ ,  $p < 0.001$ ).

### How are raters identifying high-agreement Sharing and Checking looks?

To understand the cues that raters may be relying on to make their judgements, we compared behaviour of the infant and mother coded from 11 videos that most raters agreed were Sharing Looks to the 9 videos that most raters agreed were Checking Looks (from the binomial analysis). We excluded the 3 videos where most raters agreed the infant look was an Orienting look, due to the low number. We examined whether high agreement Sharing and Checking looks differed in terms of (a) the duration of look from the infant to their mother’s face; (b) the presence of mutual gaze between mother and infant (e.g. the mother reciprocating the infant’s look to her face); (c) the presence/absence of (i) facial expressions, (ii) gestures, and (iii) vocal communication from infant to mother during the infant’s look to their mother; (d) frequency of any communicative signals produced by the infant in whole video; (e) the presence/absence of (i) facial expressions, (ii) gestures, and (iii) vocal communication from mother to the infant during the infant’s look to their mother; (f) frequency of any communicative signals produced by the mother in whole video.

**(a) the duration of look from the infant to their mother’s face.** A GLMM, with “Duration of infant look to mother” as the response variable (gamma distribution) and “Type of agreed look” (Sharing / Checking) as the independent variable and “Video Dyad ID” as a random factor, confirmed that looks judged as ‘sharing’ were significantly longer than looks judged as ‘checking’ looks ( $b = 0.84$ , SE = 0.31,  $p = 0.007$ ; Fig 1).

**(b) the duration of mutual gaze between mother and infant.** There were only 9 videos with mutual gaze (7 judged to be sharing looks; 2 judged to be checking looks), so the sample of mutual gaze occurrences was too small for inferential statistics to be appropriate. Descriptively, when mutual gaze was present it seemed to be for a longer duration in looks judged as ‘sharing’ (mean = 2.31s; SE = 0.52) than looks judged as ‘checking’ (mean = 0.53s; SE = 0.05).

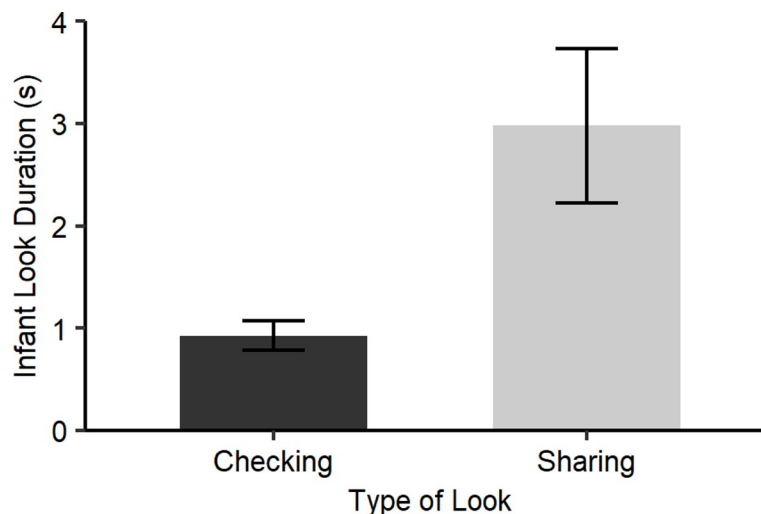
Table 3. Distribution of participant responses over the three categories of looks for the 23 looks identified as 'high agreement' by binomial tests.

Set	Mother's Answer	Number of Participants who selected 'Sharing'	Number of Participants who selected 'Checking'	Number of Participants who selected 'Orienting'
Set 1	Sharing (+)	<u>24</u> , $p < 0.001$	2	6
Set 1	Checking (-)	<u>21</u> , $p < 0.001$	5	5
Set 1	Orienting (-)	<u>21</u> , $p < 0.001$	11	0
Set 1	Sharing (+)	<u>29</u> , $p < 0.001$	0	2
Set 1	Unknown	<u>26</u> , $p < 0.001$	4	2
Set 1	Checking (+)	1	<u>23</u> , $p < 0.001$	6
Set 1	Unknown	0	4	<u>27</u> , $p < 0.001$
Set 1	Checking (-)	2	7	<u>23</u> , $p < 0.001$
Set 1	Sharing (-)	3	<u>20</u> , $p < 0.001$	9
Set 2	–	<u>21</u> , $p < 0.001$	3	8
Set 2	–	<u>20</u> , $p < 0.001$	4	8
Set 2	–	<u>30</u> , $p < 0.001$	0	1
Set 2	–	3	<u>28</u> , $p < 0.001$	0
Set 2	–	5	<u>23</u> , $p < 0.001$	2
Set 2	–	1	<u>26</u> , $p < 0.001$	3
Set 2	–	3	<u>29</u> , $p < 0.001$	0
Set 2	–	<u>21</u> , $p < 0.001$	7	4
Set 2	–	<u>21</u> , $p < 0.001$	2	8
Set 2	–	2	<u>27</u> , $p < 0.001$	3
Set 2	–	0	10	<u>22</u> , $p < 0.001$
Set 2	–	<u>22</u> , $p < 0.001$	7	3
Set 2	–	1	<u>20</u> , $p < 0.001$	10
Set 2	–	2	<u>29</u> , $p < 0.001$	1

For new video trials, the mothers' answers are recorded as well as whether they agreed (+), or disagreed (-) with the raters, or whether their ratings were unknown. The type of look selected by the highest number of participants has been underlined.

<https://doi.org/10.1371/journal.pone.0255241.t003>

(c) the presence/absence of (i) facial expressions, (ii) gestures, and (iii) vocal communication produced by the infant during their look to their mother. We then assessed whether



**Fig 1. Mean duration of look from infant to mother for 'high agreement' looks that were rated as "Checking looks" and "Sharing looks" by the majority of participants, with SEM error bars.**

<https://doi.org/10.1371/journal.pone.0255241.g001>

the presence or absence of infant facial expressions, gestures, or vocal communication during an infant's look to their mother differed for high agreement Sharing and Checking looks. We ran three GLMMs with "Presence/Absence of [communication type]" as the response variable (binomial distribution), and the independent variable was "Type of agreed look" (Sharing / Checking) with "Video Dyad ID" as a random factor (Table 4). Although descriptively more communicative signals were produced by the infant during Sharing looks than Checking looks, the models indicated that there was no significant difference in the likelihood of infant communicative signals occurring in looks judged to be 'Sharing' and 'Checking'.

**(d) the presence/absence of (i) facial expressions, (ii) gestures, and (iii) vocal communication produced by the mother during the infant's look to their mother.** We then assessed whether the presence or absence of mother facial expressions, gestures, or vocal communication during an infant's look to their mother differed for high agreement Sharing and Checking looks. There was zero or highly limited variation in at least one of the categories of looks for each of the communication behaviours (Table 5), rendering inferential statistics inappropriate. Descriptively, for all three types of communication, there is almost no mother communication present for high agreement Checking looks, while there is more communication present in high agreement Sharing looks, with 100% of these having vocal communication present.

### Are raters responding to "charismatic" (i.e. more communicative) mothers or infants?

We also wanted to test whether raters attended to the overall rates of communication from the mothers and infants in the whole video clip, and perhaps used the heuristic that more

**Table 4. The proportion of high agreement Sharing and Checking looks (as judged by the raters) in which the infant produced at least one facial expression, gesture, and vocalisation during the look to their mother, and results from the GLMM analyses.**

	Sharing	Checking	GLMM for Presence/Absence of communication type
Facial Expression Present	8/11 (0.73)	3/9 (0.33)	$b = 2.17$ , $SE = 1.80$ , $p = 0.229$
Gesture Present	4/11 (0.36)	1/9 (0.11)	$b = 8.83$ , $SE = 6.34$ , $p = 0.164$
Vocal Communication Present	7/11 (0.64)	5/9 (0.56)	$b = 0.24$ , $SE = 0.92$ , $p = 0.714$

<https://doi.org/10.1371/journal.pone.0255241.t004>

**Table 5. The proportion of high agreement Sharing and Checking looks (as judged by the raters) in which the mother produced at least one facial expression, gesture, or vocal communication during their infant's look to them.**

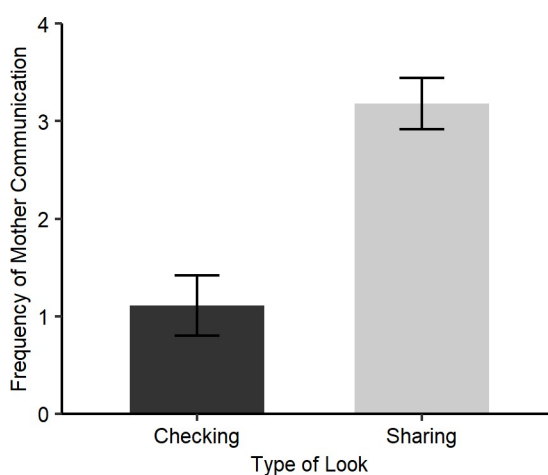
	Sharing	Checking
Facial Expression Present	4/11 (0.36)	0/9 (0.00)
Gesture Present	6/11 (0.55)	1/9 (0.11)
Vocal Communication Present	11/11 (1.00)	0/9 (0.00)

<https://doi.org/10.1371/journal.pone.0255241.t005>

communicative / charismatic individuals were more likely to share looks. To test this idea, we first ran a GLMM with “Frequency of mother communication” as the response variable (Gaussian distribution), “Type of agreed look” (Sharing / Checking) as the independent variable and “Video Dyad ID” as a random factor. Mother communication, in terms of the total number of facial expressions, gestures, and vocalisations combined, was significantly more frequent in high agreement videos rated as Sharing looks than those rated as Checking looks ( $b = 2.06$ ,  $SE = 0.43$ ,  $p < 0.001$ ; Fig 2). When we ran the same GLMM but with “Frequency of infant communication” as the response variable, we did not see the same pattern, rather infant communication occurs at similar frequencies throughout the high agreement videos rated as Checking (mean = 1.667,  $SE = 0.333$ ) and Sharing looks (mean = 2.545,  $SE = 0.312$ ;  $b = 0.809$ ,  $SE = 0.47$ ,  $p = 0.100$ ).

## Discussion

As joint attention is so pivotal in human infant development and potentially also in human evolution, our main objective in this study was to determine whether the quality of an infant's look to their mother can be reliably identified and therefore is sufficient for distinguishing whether joint attention has occurred. We found overall low agreement among naïve raters in assigning looks from infants to their mothers as Sharing, Checking, or Orienting looks, suggesting that the definitions given by Hobson and Hobson were not adequate for our raters to be able to reliably assign these types of looks in infants [47]. Whilst both our study and Hobson and Hobson's study used naïve raters to identify looks from videos of dyadic interactions, there are several key differences between our studies which may explain our failure to replicate



**Fig 2. Mean frequency of mother communication in high agreement looks rated by the majority of participants as “Checking looks” and “Sharing looks”, with SEM error bars.**

<https://doi.org/10.1371/journal.pone.0255241.g002>

their findings. Firstly, Hobson and Hobson used footage of 11-year-olds with and without autism interacting with an experimenter. Whilst they report that “mostly it was straightforward to ‘feel’ (and judge) whether the looks seen on videotape were sharing, expressive of interpersonal engagement; checking, indicative of glancing up to check the tester’s face for a reaction or instruction; or orientating” (p. 419), looks from preverbal infants proved much more difficult to judge for our raters. Cues as to the intention underlying looks in preverbal infants may be subtler and more difficult to distinguish than in the older children in Hobson and Hobson’s study.

There are also other important methodological differences between the studies: our raters were constrained to viewing a short video clip once at normal speed, whereas Hobson and Hobson don’t specify any constraints on how much video surrounding the looks was reviewed, how many times raters viewed the videos, or at what speed. Second, whereas Hobson and Hobson had 2 raters assess just 27 looks from 6 children, we had 30 raters assess 60 looks from 15 infants, providing a more representative sample of both raters and looks. Whilst these differences mean that our study does not challenge Hobson and Hobson’s original findings, our results question whether they can be extended to preverbal infants. In our controlled experiment, Sharing looks could not be reliably identified by naïve third-party observers in preverbal infants, suggesting that types of look should not be considered as a valid way of identifying joint attention events in preverbal infants.

We found not only that naïve observers had low levels of agreement with each other, but critically they rarely agreed with the judgement of the mother who received the looks. When considering the 26 looks that mothers categorised as sharing, checking or orienting, in only 3/26 cases did a significant majority of raters agree with the mothers’ judgement of the type of look they received from their infants. It is possible that when engaged in an interaction and receiving a look directly, an individual can accurately infer the intentions of their interaction partner, and we hoped that the mothers’ judgements made immediately after the play interaction would be an approximation of this experience. Our data shows clearly, however, that third party observers seem unable to access the same cues that the mothers received directly from their infants. This is perhaps not surprising as joint attention arises when partners dynamically perceive, interpret, and respond to behaviours in each other. As researchers, we cannot access a mother’s experience of being involved in an interaction with her infant. Perhaps future research could assess if types of look can be reliably identified by multiple observers from footage obtained from head mounted cameras on the interaction partners. Whilst this doesn’t replicate the experience of being the interaction partner, it may give a better approximation of the mothers’ perspective of their interaction with their infant whilst still allowing the reliability of judgements to be assessed, which is fundamental to replicable, reliable scientific investigation.

We considered that people with more experience with infants may be more reliable at assigning the types of look (i.e. have higher levels of agreement), so we examined the responses of mothers and non-mothers separately. However, we found that mothers and non-mothers showed similar low levels of agreement on the type of look. Experience with young infants does not appear to make raters more reliable in coding Sharing, Checking, and Orienting looks. Another interesting finding was that raters were significantly more confident for high agreement looks compared to low agreement looks. We don’t know the extent to which participants explicitly or implicitly relied on the behavioural cues that we coded in the videos, but given the positive relationship between a rater’s self-report of confidence in their judgement and agreement between raters, future research may be able to usefully ask raters to reflect on the cues they relied on.

Taken together, our results indicate that in our study naïve third party observers were unable to categorise most infant looks reliably. Future research could investigate if reliability

can be improved by changing parameters of the current study, such as raters receiving more training before completing the task or raters watching longer clips that contain more contextual information, more times or in slow motion. Until robust evidence of parameters that may support reliable identification of look type by 3<sup>rd</sup> party observers is available, we suggest that type of look should not be used to assess whether joint attention events have occurred in infants. Hobson and Hobson was the only paper we found that made explicit claims about the quality of looks as valid markers of joint attention [47], but there are many papers that implicitly use quality of looks to determine whether children “acknowledge their partner’s participation” (Table 1), and subsequently whether joint attention has occurred. While the frequency and duration of looks between partners can be readily and objectively extracted from an interaction through video coding, our low levels of agreement on the types of look suggest that rating the quality of looks could be an unreliable way of assessing whether partners understand that they are both attending to the same object/event together and therefore engaging in a joint attention event.

We identified a subset of looks where significantly more raters than expected by chance agreed on the type of look as ‘Checking’ or ‘Sharing’ (only 3 looks were agreed to be Orienting looks, which was too few to analyse). Video coding of infant and mother behaviour during these high agreement videos revealed several potential behaviour cues that raters may have been relying on to distinguish these types of look. ‘Sharing’ looks were significantly longer than ‘Checking’ looks and, descriptively, were more likely to contain mutual gaze, and when it did occur, mutual gaze periods seemed to be longer. Gaze, attention shifts, gaze alternation, and mutual gaze are already used across many definitions of joint attention events (Table 1), and so including a measure of duration for looks and mutual gaze would be an objective and readily code-able variable to examine in future studies on joint attention.

When we looked at communication, there were no significant effects of infant communication on how the raters assigned looks. Because we had instructed the raters to attend to the infant’s look to the mother, we expected that they would rely more on the infant’s behaviour than the mother’s. Surprisingly, we found the mother’s but not the infant’s overall rates of communication varied significantly with the classification of high agreement ‘Sharing’ and ‘Checking’ looks, indicating this may have been a cue the raters relied on to make their judgements. Taken together, it is possible that in the minority of cases where raters agreed at above chance rates, they may have been operating with the heuristics that longer or mutual looks are indicative of a sharing intent and that more communicative mothers are more likely to receive sharing looks from their infant. Whilst duration of look is a relevant cue, using the mother’s communicativeness to make judgements about the infant’s intentions is unlikely to be a reliable cue. This highlights the limitations of asking raters to make holistic judgements of stimuli: it is impossible to know which behavioural cues raters relied on at either an explicit or implicit level, and if raters are relying upon irrelevant cues their judgements would be invalid. We therefore advocate focussing on objective behavioural criteria for assessing sharing intentions, rather than rating subjective characteristics of behaviour.

Our study was conducted in a single species and in a single cultural context, providing favourable conditions for reliable judgements to be made, but our raters still failed to agree on the majority of their judgements. If we are to fully understand the ontogeny of joint attention we need to address the persistent sampling bias in developmental psychology for W.E.I.R.D. (Western, Educated, Industrialized, Rich, and Democratic) samples [113], and study early joint attention in diverse cultural contexts. Equally, in order to understand the evolutionary origins of joint attention we need to examine interactions in other species. As our results cast doubt on whether types of look can be reliably determined within a single species and cultural context, it seems highly unlikely they could be reliably identified in cross-cultural or cross-



species comparisons. Whilst subjective holistic judgements of behavioural intentions can seem appealing in that they may have the potential to capture nuanced, subtle cues that quantitative measurement of behaviour via video coding may miss, our study indicates that we need to move away from subjective approaches that assess the quality of looks and engagement based on a holistic viewing of an interaction. Moving towards an approach where the 'jointness' of a joint attention event and awareness of attending to the object/event together is based on extraction of detailed behaviours from videos, such as look duration, mutual gaze duration, and communication by both interactants, will yield conservative, but importantly reliable identification of joint attention events that can be applied across cultures and species. Looking ahead, the implementation of rigorous, reliable measures of behaviour will be crucial for understanding the ontogenetic and evolutionary origins of joint attention.

## Supporting information

**S1 Table.** Confusion matrix between how mothers assigned their infants' looks and how raters assigned the looks for which they had high agreement.

(DOCX)

**S1 Data.**

(XLSX)

**S1 File.**

(TXT)

## Acknowledgments

Many thanks to the mothers and infants who participated in this study; to Junior Whiteley who assisted in collecting and cutting the videos used; and to Yujin Lee for help with early versions of the video coding scheme.

## Author Contributions

**Conceptualization:** Kirsty E. Graham, Joanna C. Bury-Weitzel, Nicole J. Lahiff, Claudia Wilke, Katie E. Slocombe.

**Data curation:** Kirsty E. Graham, Claudia Wilke.

**Formal analysis:** Kirsty E. Graham, Claudia Wilke.

**Funding acquisition:** Katie E. Slocombe.

**Investigation:** Kirsty E. Graham, Joanna C. Bury-Weitzel, Nicole J. Lahiff, Katie E. Slocombe.

**Methodology:** Kirsty E. Graham, Claudia Wilke, Katie E. Slocombe.

**Project administration:** Kirsty E. Graham, Katie E. Slocombe.

**Resources:** Katie E. Slocombe.

**Software:** Katie E. Slocombe.

**Supervision:** Katie E. Slocombe.

**Validation:** Claudia Wilke.

**Visualization:** Kirsty E. Graham.

**Writing – original draft:** Kirsty E. Graham, Joanna C. Buryan-Weitzel, Nicole J. Lahiff, Claudia Wilke, Katie E. Slocombe.

**Writing – review & editing:** Kirsty E. Graham, Joanna C. Buryan-Weitzel, Nicole J. Lahiff, Claudia Wilke, Katie E. Slocombe.

## References

1. Moore C, Dunham PJ, editors. Joint Attention: Its Origins and Role in Development. Mahwah, NJ: Lawrence Erlbaum Associates; 1995.
2. Bakeman R, Adamson LB. Coordinating Attention to People and Objects in Mother-Infant and Peer-Infant Interaction. *Child Dev.* 1984; 55(4):1278–89. PMID: [6488956](#)
3. Carpenter M, Nagell K, Tomasello M, Butterworth G, Moore C. Social Cognition, Joint attention, and Communicative Competence from 9 to 15 Months of Age. *Monogr Soc Res Child Dev.* 1998; 63(4):1–174.
4. Adamson LB, Bakeman R, Suma K, Robins DL. An Expanded View of Joint Attention: Skill, Engagement, and Language in Typical Development and Autism. *Child Dev.* 2019; 90(1):e1–18. <https://doi.org/10.1111/cdev.12973> PMID: [28991358](#)
5. Tomasello M, Farrar MJ. Joint Attention and Early Language. *Child Dev.* 1986; 57(6):1454–63. PMID: [3802971](#)
6. Tomasello M, Carpenter M, Call J, Behne T, Moll H. Understanding and sharing intentions: The origins of cultural cognition. *Behav Brain Sci.* 2005; 28(5):675–91. <https://doi.org/10.1017/S0140525X05000129> PMID: [16262930](#)
7. Carpenter M, Tomasello M. Joint Attention and Imitative Learning in Children, Chimpanzees, and Enculturated Chimpanzees. *Soc Dev.* 1995; 4(3):217–37.
8. Bard KA, Leavens DA. Socioemotional factors in the development of joint attention in human and ape infants. In: Roeska-Hardy, Neumann-Held, editors. *Learning from Animals?: Examining the Nature of Human Uniqueness.* New York: Psychology Press Ltd; 2009. p. 89–104.
9. Pitman CA, Shumaker RW. Does Early Care Affect Joint Attention in Great Apes (Pan troglodytes, Pan paniscus, Pongo abelii, Pongo pygmaeus, Gorilla gorilla)? *J Comp Psychol.* 2009; 123(3):334–41. <https://doi.org/10.1037/a0015840> PMID: [19685976](#)
10. Bard KA, Bakeman R, Boysen ST, Leavens DA. Emotional engagements predict and enhance social cognition in young chimpanzees. *Dev Sci.* 2014; 17(5):682–96. <https://doi.org/10.1111/desc.12145> PMID: [24410843](#)
11. Siposova B, Carpenter M. A new look at joint attention and common knowledge. *Cognition.* 2019; 189:260–74. <https://doi.org/10.1016/j.cognition.2019.03.019> PMID: [31015079](#)
12. Aarne P, Tallberg IM. Visual check back in children with specific language impairment. *J Pragmat.* 2010; 42(11):3106–3113.
13. Adamson LB, Bakeman R, Deckner DF, Ronski M. Joint engagement and the emergence of language in children with autism and down syndrome. *J Autism Dev Disord.* 2009; 39(1):84–96. <https://doi.org/10.1007/s10803-008-0601-7> PMID: [18581223](#)
14. Adamson LB, Deckner DF, Bakeman R. Early interests and joint engagement in typical development, autism, and down syndrome. *J Autism Dev Disord.* 2010; 40(6):665–676. <https://doi.org/10.1007/s10803-009-0914-1> PMID: [20012678](#)
15. Adamson LB, Ronski M, Bakeman R, Sevcik RA. Augmented language intervention and the emergence of symbol-infused joint engagement. *J Speech Lang Hear Res.* 2010; 53(6):1769–1773. [https://doi.org/10.1044/1092-4388\(2010/09-0208\)](https://doi.org/10.1044/1092-4388(2010/09-0208)) PMID: [20705741](#)
16. Aldrich NJ, Brooks PJ, Yuksel-Sokmen PO, Ragir S, Flory MJ, Lennon EM, et al. Infant twins' social interactions with caregivers and same-age siblings. *Infant Behav Dev.* 2015; 41:127–141. <https://doi.org/10.1016/j.infbeh.2015.08.005> PMID: [26476957](#)
17. Allely CS, Johnson PC, Marwick H, Lidstone E., Kočovská E, Puckering C, et al. Prediction of 7-year psychopathology from mother-infant joint attention behaviours: a nested case–control study. *BMC Pediatr.* 2013; 13(1):147. <https://doi.org/10.1186/1471-2431-13-147> PMID: [24063312](#)
18. Arens K, Cress CJ, Marvin CA. Gaze-shift patterns of young children with developmental disabilities who are at risk for being nonspeaking. *Educ Train Dev Disabil.* 2005; 40(2):158–170.
19. Arnold A, Semple RJ, Beale I, Fletcher-Flinn CM. Eye contact in children's social interactions: What is normal behaviour? *J Intellect Dev Disabil.* 2000; 25(3):207–216.

20. Benigno JP, Clark L, Farrar MJ. Three is not always a crowd: Contexts of joint attention and language. *J Child Lang*. 2007; 34(1):175–187. <https://doi.org/10.1017/s0305000906007732> PMID: 17340943
21. Benigno JP, Farrar MJ. Determinants of joint attention in young siblings' play. *Infant Child Dev*. 2012; 21(2):160–174.
22. Bigelow AE. The development of joint attention in blind infants. *Dev Psychopathol*. 2003; 15(2):259–275. <https://doi.org/10.1017/s0954579403000142> PMID: 12931827
23. Bigelow AE, MacLean K, Proctor J. The role of joint attention in the development of infants' play with objects. *Dev Sci*. 2004; 7(5):518–526. <https://doi.org/10.1111/j.1467-7687.2004.00375.x> PMID: 15603284
24. Bigelow AE, MacLean K, Proctor J, Myatt T, Gillis R, Power M. Maternal sensitivity throughout infancy: Continuity and relation to attachment security. *Infant Behav Dev*. 2010; 33(1):50–60. <https://doi.org/10.1016/j.infbeh.2009.10.009> PMID: 20004022
25. Brune CW, Woodward AL. Social cognition and social responsiveness in 10-month-old infants. *J Cogn Dev*. 2007; 8(2):133–158.
26. Charman T, Baron-Cohen S, Swettenham J, Baird G, Cox A, Drew A. Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cogn Dev*. 2000; 15(4):481–498.
27. Charman T, Baron-Cohen S, Swettenham J, Baird G, Drew A, Cox A. Predicting language outcome in infants with autism and pervasive developmental disorder. *Int J Lang Commun Disord*. 2003; 38(3):265–285. <https://doi.org/10.1080/13682031000104830> PMID: 12851079
28. Chiang C-H, Chu C-L, Lee T-C. Efficacy of caregiver-mediated joint engagement intervention for young children with autism spectrum disorders. *Autism*. 2016; 20(2):172–182. <https://doi.org/10.1177/1362361315575725> PMID: 25896268
29. Childers JB, Vaughan J, Burquest DA. Joint attention and word learning in Ngas-speaking toddlers in Nigeria. *J Child Lang*. 2007; 34(2):199–225. <https://doi.org/10.1017/s0305000906007835> PMID: 17542156
30. Christidou D. Art on the move: The role of joint attention in visitors' encounters with artworks. *Learn Cult Soc Interact*. 2018; 19:1–10.
31. Cleveland A, Striano T. The effects of joint attention on object processing in 4- and 9-month-old infants. *Infant Behav Dev*. 2007; 30(3):499–504. <https://doi.org/10.1016/j.infbeh.2006.10.009> PMID: 17610957
32. Clifford SM, Dissanayake C. The early development of joint attention in infants with autistic disorder using home video observations and parental interview. *J Autism Dev Disord*. 2008; 38(5):791–805. <https://doi.org/10.1007/s10803-007-0444-7> PMID: 17917803
33. Clifford S, Dissanayake C. Dyadic and triadic behaviours in infancy as precursors to later social responsiveness in young children with autistic disorder. *J Autism Dev Disord*. 2009; 39(10):1369–1380. <https://doi.org/10.1007/s10803-009-0748-x> PMID: 19475503
34. Conboy BT, Brooks R, Meltzoff AN, Kuhl PK. Social interaction in infants' learning of second-language phonetics: An exploration of brain-behavior relations. *Dev Neuropsychol*. 2015; 40(4):216–229. <https://doi.org/10.1080/87565641.2015.1014487> PMID: 26179488
35. Craig TKJ, Bialas I, Hodson S, Cox AD. Intergenerational transmission of somatization behaviour: 2. Observations of joint attention and bids for attention. *Psychol Med*. 2004; 34(2):199. <https://doi.org/10.1017/s0033291703001120> PMID: 14982126
36. Cress C, Arens K, Zajicek A. Comparison of Engagement Patterns of Young Children with Developmental Disabilities between Structured and Free Play. *Educ Train Dev Disabil*. 2007; 42(2):152–164.
37. de la Ossa JL, Gauvain M. Joint attention by mothers and children while using plans. *Int J Behav Dev*. 2001; 25(2):176–183.
38. Depowski N, Abaya H, Oghalai J, Bortfeld H. Modality use in joint attention between hearing parents and deaf children. *Front Psychol*. 2015; 6:1556. <https://doi.org/10.3389/fpsyg.2015.01556> PMID: 26528214
39. Gaffan EA, Matrins C, Healy S, Murray L. Early social experience and individual differences in infants' joint attention. *Soc Dev*. 2010; 19(2):369–393.
40. Gauthier K, Genesee F, Dubois ME, Kasparian K. Communication patterns between internationally adopted children and their mothers: Implications for language development. *Appl Psycholinguist*. 2011; 34:337–359.
41. Gonsiorowski A, Williamson RA, Robin DL. Imitation of object-directed acts in young children with autism spectrum disorders. *J Autism Dev Disord*. 2016; 46(2):691–697. <https://doi.org/10.1007/s10803-015-2596-1> PMID: 26386710
42. Grossmann T, Johnson MH. Selective prefrontal cortex responses to joint attention in early infancy. *Biol Lett*. 2010; 6(4):540–543. <https://doi.org/10.1098/rsbl.2009.1069> PMID: 20106861

43. Gulsrud AC, Helleman G, Shire S, Kasari C. Isolating active ingredients in a parent-mediated social communication intervention for toddlers with autism spectrum disorder. *J Child Psychol Psychiatry*. 2016; 57(5):606–613. <https://doi.org/10.1111/jcpp.12481> PMID: 26525461
44. Gulsrud AC, Kasari C, Freeman S, Paparella T. Children with autism's response to novel stimuli while participating in interventions targeting joint attention or symbolic play skills. *Autism*. 2007; 11(6):535–546. <https://doi.org/10.1177/1362361307083255> PMID: 17947289
45. Hahn LJ, Brady NC, Fleming KK, Warren SF. Joint engagement and early language in young children with fragile X syndrome. *J Speech Lang Hear Res*. 2016; 59(5):1087–1098. [https://doi.org/10.1044/2016\\_JSLHR-L-15-0005](https://doi.org/10.1044/2016_JSLHR-L-15-0005) PMID: 27681878
46. Henderson EN, Jennings KD. Maternal depression and the ability to facilitate joint attention with 18-month-olds. *Infancy*. 2003; 4(1):27–46.
47. Hobson JA, Hobson RP. Identification: The missing link between joint attention and imitation? *Dev Psychopathol*. 2007; 19:411–431. <https://doi.org/10.1017/S0954579407070204> PMID: 17459177
48. Hughes ME, Allen SE. The effect of individual discourse-pragmatic features on referential choice in child English. *J Pragmat*. 2013; 56:15–30.
49. Ine H, Heleen N, Bea M. Attentional processes in interactions between people with profound intellectual and multiple disabilities and direct support staff. *Res Dev Disabil*. 2011; 32(2):491–503. <https://doi.org/10.1016/j.ridd.2010.12.034> PMID: 21236632
50. Ingersoll B, Schreibman L. Teaching reciprocal imitation skills to young children with autism using a naturalistic behavioral approach: Effects on language, pretend play, and joint attention. *J Autism Dev Disord*. 2006; 36(4):487–505. <https://doi.org/10.1007/s10803-006-0089-y> PMID: 16568355
51. Johnson KC, DesJarding JL, Barker DH, Quittner AL, Winter ME. Assessing joint attention and symbolic play in children with cochlear implants and multiple disabilities: Two case studies. *Otol Neurotol*. 2008; 29(2):246–250. <https://doi.org/10.1097/mao.0b013e318162f1f3> PMID: 18223452
52. Kasari C, Gulsrud AC, Wong C, Kwon S, Locke J. Randomized controlled caregiver mediated joint engagement intervention for toddlers with autism. *J Autism Dev Disord*. 2010; 40(9):1045–1056. <https://doi.org/10.1007/s10803-010-0955-5> PMID: 20145986
53. Ketelaar L, Rieffe C, Wieferrink CH, Frijns JH. Does hearing lead to understanding? Theory of mind in toddlers and preschoolers with cochlear implants. *J Pediatr Psychol*. 2012; 37(9):1041–1050. <https://doi.org/10.1093/jpepsy/jss086> PMID: 22847880
54. Kompatsiari K, Ciardo F, Tikhanoff V, Metta G, Wykowska A. On the role of eye contact in gaze cueing. *Sci Rep*. 2018; 8(1):1–10. <https://doi.org/10.1038/s41598-017-17765-5> PMID: 29311619
55. Landry S, Smith K, Swank P, Guttentag C. A responsive parenting intervention: The optimal timing across early childhood for impacting maternal behaviors and child outcomes. *Dev Psychol*. 2008; 44(5):1335–1353. <https://doi.org/10.1037/a0013030> PMID: 18793067
56. Larkin F, Guerin S, Hobson JA, Gutstein SE. The Relationship Development Assessment—Research Version: Preliminary validation of a clinical tool and coding schemes to measure parent-child interaction in autism. *Clin Child Psychol Psychiatry*. 2015; 20(2):239–260. <https://doi.org/10.1177/1359104513514065> PMID: 24366957
57. Legerstee M, Markova G, Fisher T. The role of maternal affect attunement in dyadic and triadic communication. *Infant Behav Dev*. 2007; 30(2):296–306. <https://doi.org/10.1016/j.infbeh.2006.10.003> PMID: 17400046
58. Macdonald RG, Tatler BW. Gaze in a real-world social interaction: A dual eye-tracking study. *Q J Exp Psychol*. 2018; 71(10):2162–2173.
59. Markus J, Mundy P, Morales M, Delgado CE, Yale M. Individual differences in infant skills as predictors of child-caregiver joint attention and language. *Soc Dev*. 2000; 9(3):302–315.
60. Mastin JD, Vogt P. Infant engagement and early vocabulary development: a naturalistic observation study of Mozambican infants from 1;1 to 2;1. *J Child Lang*. 2016; 43(2):235–264. <https://doi.org/10.1017/S0305000915000148> PMID: 26087953
61. Matatyaho DJ, Gogate LJ. Type of maternal object motion during synchronous naming predicts pre-verbal infants' learning of word-object relations. *Infancy*. 2008; 13(2):172–184. <https://doi.org/10.1080/15250000701795655> PMID: 33412725
62. Morales M, Mundy P, Crowson M, Neal AR, Delgado C. Individual differences in infant attention skills, joint attention, and emotion regulation behaviour. *Int J Behav Dev*. 2005; 29(3):259–263.
63. Morgan B, Maybery M, Durkin K. Weak central coherence, poor joint attention, and low verbal ability: Independent deficits in early autism. *Dev Psychol*. 2003; 39(4):646. <https://doi.org/10.1037/0012-1649.39.4.646> PMID: 12859119

64. Najnin S, Banerjee B. (2018). Pragmatically Framed Cross-Situational Noun Learning Using Computational Reinforcement Models. *Front Psychol.* 2018; 9(5):1–18. <https://doi.org/10.3389/fpsyg.2018.00005> PMID: [29441027](https://pubmed.ncbi.nlm.nih.gov/29441027/)
65. Neerinx H, Vos P, Van Den Noortgate W, Maes B. Temporal analysis of attentional processes in spontaneous interactions between people with profound intellectual and multiple disabilities and their support workers. *J Intellect Disabil Res.* 2014; 58(8):721–733. <https://doi.org/10.1111/jir.12067> PMID: [23870049](https://pubmed.ncbi.nlm.nih.gov/23870049/)
66. Nelson PB, Adamson LB, Bakeman R. Toddlers' joint engagement experience facilitates preschoolers' acquisition of theory of mind. *Dev Sci.* 2008; 11(6):847–852. <https://doi.org/10.1111/j.1467-7687.2008.00733.x> PMID: [19046153](https://pubmed.ncbi.nlm.nih.gov/19046153/)
67. Nordahl-Hansen A, Fletcher-Watson S, McConachie H, Kaale A. Relations between specific and global outcome measures in a social-communication intervention for children with autism spectrum disorder. *Res Autism Spectr Disord.* 2016;29–30:19–29.
68. Nowakowski ME, Tasker SL, Cunningham CE, McHolm AE, Edison S, Pierre JS, et al. Joint attention in parent–child dyads involving children with selective mutism: A comparison between anxious and typically developing children. *Child Psychiatry Hum Dev.* 2011; 42(1):78–92. <https://doi.org/10.1007/s10578-010-0208-z> PMID: [20960051](https://pubmed.ncbi.nlm.nih.gov/20960051/)
69. Nowakowski ME, Tasker SL, Schmidt LA. Establishment of joint attention in dyads involving hearing mothers of deaf and hearing children, and its relation to adaptive social behavior. *Am Ann Deaf.* 2009; 154(1):15–29. <https://doi.org/10.1353/aad.0.0071> PMID: [19569301](https://pubmed.ncbi.nlm.nih.gov/19569301/)
70. Nowakowski ME, Tasker SL, Schmidt LA. Joint attention in Toddlerhood Predicts internalizing problems at Early School age. *Clin Pediatr.* 2012; 51(11):1032–1040. <https://doi.org/10.1177/0009922812441670> PMID: [22511189](https://pubmed.ncbi.nlm.nih.gov/22511189/)
71. Pierce L, Genesee F, Gauthier K, Dubois ME. Communication patterns between parents and children: Comparing mothers and fathers in different learner contexts. *Appl Psycholinguist.* 2015; 36(5):1223.
72. Povis KT, Crowley K. Family learning in object-based museums: The role of joint attention. *Visitor Studies.* 2015; 18(2):168–182.
73. Rice DE, Adamson LB, Winner E, McGee GG. (2016). A cross-sectional study of shared attention by children with autism and typically developing children in an inclusive preschool setting. *Top Lang Disord.* 2016; 36(3):245–265.
74. Richardson D, Dale R, Kirkham N. The Art of Conversation Is Coordination: Common Ground and the Coupling of Eye Movements during Dialogue. *Psychol Sci.* 2007; 18(5):407–413. <https://doi.org/10.1111/j.1467-9280.2007.01914.x> PMID: [17576280](https://pubmed.ncbi.nlm.nih.gov/17576280/)
75. Sanefuji W, Yamashita H, Ohgami H. Shared minds: Effects of a mother's imitation of her child on the mother–child interaction. *Infant Ment Health J.* 2009; 30(2):145–157. <https://doi.org/10.1002/imhj.20208> PMID: [28636174](https://pubmed.ncbi.nlm.nih.gov/28636174/)
76. Saxon TF, Colombo J, Robinson EL, Frick JE. Dyadic interaction profiles in infancy and preschool intelligence. *J Sch Psychol.* 2000; 38(1):9–25.
77. Schechter DS, Willheim E, Hinojosa C, Scholfield-Kleinman K, Turner JB, McCaw J, et al. Subjective and objective measures of parent-child relationship dysfunction, child separation distress, and joint attention. *Psychiatry.* 2010; 73(2):130–144. <https://doi.org/10.1521/psyc.2010.73.2.130> PMID: [20557225](https://pubmed.ncbi.nlm.nih.gov/20557225/)
78. Schertz HH, Call-Cummings M, Horn K, Quest K, Law RS. Social and instrumental interaction between parents and their toddlers with autism: A qualitative analysis. *J Early Interv.* 2018; 40(1):20–38.
79. Schertz HH, Odom SL. Joint attention and early intervention with autism: A conceptual framework and promising approaches. *J Early Interv.* 2004; 27(1):42–54.
80. Roy P, Chiat S. Developmental pathways of language and social communication problems in 9–11 year olds: Unpicking the heterogeneity. *Res Dev Disabil.* 2014; 35(10):2534–2546. <https://doi.org/10.1016/j.ridd.2014.06.014> PMID: [25005063](https://pubmed.ncbi.nlm.nih.gov/25005063/)
81. Shire SY, Gulsrud A, Kasari C. Increasing responsive parent–child interactions and joint engagement: Comparing the influence of parent-mediated intervention and parent psychoeducation. *J Autism Dev Disord.* 2016; 46(5):1737–1747. <https://doi.org/10.1007/s10803-016-2702-z> PMID: [26797940](https://pubmed.ncbi.nlm.nih.gov/26797940/)
82. Skarabela B. Signs of early social cognition in children's syntax: The case of joint attention in argument realization in child Inuktitut. *Lingua.* 2007; 117(11):1837–1857.
83. Skarabela B, Allen SE. How newness and joint attention work together in child Inuktitut: Assessing discourse-pragmatic models of early argument realization. In *Proceedings of the 34th Annual Boston University Conference on Language Development.* 2010 (pp. 385–396).
84. Slaughter V, Peterson CC, Carpenter M. Maternal talk about mental states and the emergence of joint visual attention. *Infancy.* 2008; 13(6):640–659.



85. Smith JL, McCarthy JW, Benigno JP. The effect of high-tech AAC system position on the joint attention of infants without disabilities. *Augment Altern Commun.* 2009; 25(3):165–175. <https://doi.org/10.1080/07434610902972410> PMID: 19639478
86. Spector V, Charlop MH. A sibling-mediated intervention for children with autism spectrum disorder: Using the natural language paradigm (NLP). *J Autism Dev Disord.* 2018; 48(5):1508–1522. <https://doi.org/10.1007/s10803-017-3404-x> PMID: 29170941
87. Striano T, & Bertin E. Social-cognitive skills between 5 and 10 months of age. *Br J Dev Psychol.* 2005; 23(4):559–568. <https://doi.org/10.1348/026151005X26282> PMID: 21214597
88. Striano T, Bertin E. Coordinated affect with mothers and strangers: A longitudinal analysis of joint engagement between 5 and 9 months of age. *Cogn Emot.* 2005; 19(5):781–790.
89. Striano T, Stah D, Clevelan A. Taking a closer look at social and cognitive skills: A weekly longitudinal assessment between 7 and 10 months of age. *Eur J Dev Psychol.* 2009; 6(5):567–591,
90. Suma K, Adamson LB, Bakeman R, Robins DL, Abrams DN. After early autism diagnosis: Changes in intervention and parent–child interaction. *J Autism Dev Disord.* 2016; 46(8):2720–2733. <https://doi.org/10.1007/s10803-016-2808-3> PMID: 27193183
91. Sung J, Hsu H-C. Korean mothers' attention regulation and referential speech: Associations with language and play in 1-year-olds. *Int J Behav Dev.* 2009; 33(5):430–439.
92. Tasker SL, Nowakowski ME, Schmidt LA. Joint attention and social competence in deaf children with cochlear implants. *J Dev Phys Disabil.* 2010; 22(5):509–532.
93. Tasker SL, Schmidt LA. The “dual usage problem” in the explanations of “joint attention” and children's socioemotional development: A reconceptualization. *Dev Rev.* 2008; 28(3):263–288.
94. Trautman CH, Rosenthal Rollins P. Child-centered behaviors of caregivers with 12-month-old infants: Associations with passive joint engagement and later language. *Appl Psycholinguist.* 2006; 27(3):447–463.
95. Vismara L, McCormick CE, Wagner AL, Monlux K, Nadhan A, Young GS. Telehealth parent training in the Early Start Denver Model: Results from a randomized controlled study. *Focus Autism Other Dev Disabil.* 2018; 33(2):67–79.
96. Walberg J, Craig-Unkefer L. An examination of the effects of a social communication intervention on the play behaviors of children with autism spectrum disorder. *Educ Train Autism Dev Disabil.* 2010; 45(1):69–80.
97. Yazbe A, D'Entremont B. A longitudinal investigation of the still-face effect at 6 months and joint attention at 12 months. *Br J Dev Psychol.* 2006; 24:589–601.
98. Yont KM, Snow CE, Vernon-Feagans L. Is chronic otitis media associated with differences in parental input at 12 months of age? An analysis of joint attention and directives. *Appl Psycholinguist.* 2003; 24(4):581.
99. Yoon S, Kelso A G, Lock A, Lyons-Ruth K. Mother–infant joint attention and sharing: Relations to disorganized attachment and maternal disrupted communication. *J Genet Psychol.* 2014; 175(6):494–510.
100. Yu C, Smith LB. Joint attention without gaze following: Human infants and their parents coordinate visual attention to objects through eye-hand coordination. *PLoS One.* 2013; 8(11):e79659. <https://doi.org/10.1371/journal.pone.0079659> PMID: 24236151
101. Yu C, Smith LB. The social origins of sustained attention in one-year-old human infants. *Curr Biol.* 2016; 26(9):1235–1240. <https://doi.org/10.1016/j.cub.2016.03.026> PMID: 27133869
102. Yu C, Smith LB. Multiple sensory-motor pathways lead to coordinated visual attention. *Cogn Sci.* 2017; 41:5–31. <https://doi.org/10.1111/cogs.12366> PMID: 27016038
103. Yu C., & Smith L. B. Hand–eye coordination predicts joint attention. *Child Dev.* 2017; 88(6):2060–2078. <https://doi.org/10.1111/cdev.12730> PMID: 28186339
104. Zampini L, Salvi A, D'Odorico L. Joint attention behaviours and vocabulary development in children with Down syndrome. *J Intellect Disabil Res.* 2015; 59(10): 891–901. <https://doi.org/10.1111/jir.12191> PMID: 25727094
105. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available from: <https://www.r-project.org/>
106. Geitater M, Lemon J, Singh I. irr: Various Coefficients of Interrater Reliability and Agreement. R package version 0.84.1. [Internet]. 2019. Available from: <https://cran.r-project.org/package=irr>
107. Bates D, Maechler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. *J Stat Softw.* 2015; 67(1):1–48.
108. Kuznetsova A, Brockhoff P, Christensen R. lmerTest Package: Tests in Linear Mixed Effects Models. *J Stat Softw.* 2017; 82(13):1–26.
109. Wickham H. The Split-Apply-Combine Strategy for Data Analysis. *J Stat Softw.* 2011; 40(1):1–29.



110. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag; 2016.
111. Fleiss JL. Measuring nominal scale agreement among many raters. Psychol Bull. 1971; 76(5):378–82.
112. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics. 1977; 33:159–74. PMID: [843571](#)
113. Nielsen M, Haun D, Kärtner J, Legare CH. The persistent sampling bias in developmental psychology: A call to action. J Exp Child Psychol. 2017; 162:31–8. <https://doi.org/10.1016/j.jecp.2017.04.017> PMID: [28575664](#)